Collaborative Virtual Environments;
An introductory review of issues and systems

Elizabeth F. Churchill
FX Palo Alto Laboratory Inc.
3400, Hillview Avenue,
Palo Alto
California
94304
USA
Email: churchill@pal.xerox.com

Dave Snowdon
Department of Computer Science,
University of Nottingham,
University Park,
Nottingham
NG9 2RD
UK
Email: d.snowdon@cs.nott.ac.uk

Abstract

A Collaborative Virtual Environment or CVE is a distributed, virtual reality that is designed to support collaborative activities. As such, CVEs provide a potentially infinite, graphically realized digital landscape within which multiple users can interact with each other and with simple or complex data representations. CVEs are increasingly being used to support collaborative work between geographically separated and between collocated collaborators. CVEs vary in the sophistication of the data and embodiment representations employed and in the level of interactivity supported. It is clear that systems intended to support individual and collaborative activities must be designed with the user's social and cognitive task characteristics and requirements considered explicitly. In this paper, we detail a number of existing systems and applications, but first discuss the nature of collaborative and cooperative work activities and consider the place of virtual reality systems in supporting such collaborative work. Following this, we discuss some future research directions.

1. INTRODUCTION

Information sharing is central to collaborative work. Collaborative virtual environments (CVEs) are distributed virtual reality systems that offer graphically realized, potentially infinite, digital landscapes. Within these landscapes, individuals can share information through interaction with each other and through individual and collaborative interaction with data representations.

Virtual environments offer the potential for much flexibility in the way the landscape(s), data and individuals are represented. For example, representations of data or of individuals may utilize sophisticated 3D graphical representations, be 2D 'flat' representations or be simply text flows pasted onto planes. Landscapes can be large or small, coloured planes, complex meshed surfaces, detailed fantastic vistas or photo-realistic backdrops. The example in Figure 1 from Benford *et al.* (1997b) illustrates this.

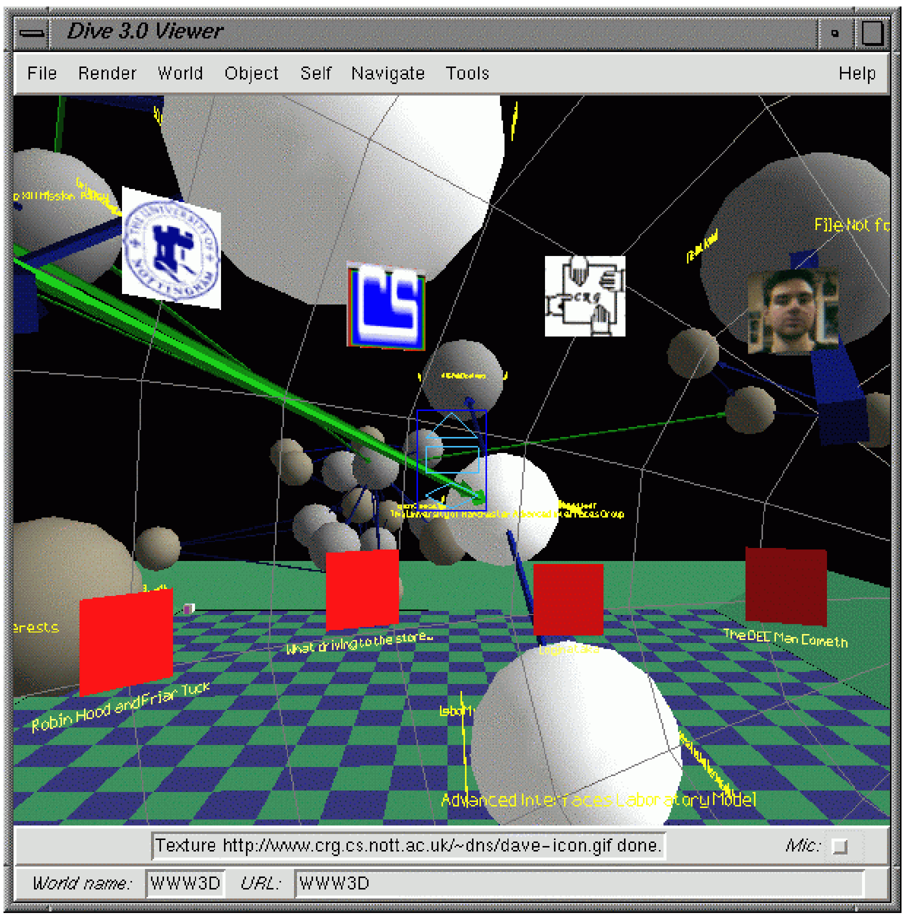

Figure 1: WWW3D: a mulit-user web browser

Figure 1 shows WWW-3D, an application for visualizing parts of the web that have been explored. In this figure a representation of part of the web is shown. In the foreground are graphical and textual elements and in the background is a checkered ground plane below a black “sky”. In addition, the background reveals other web pages, represented as opaque spheres. Users can navigate through the space and “enter” the opaque spheres, which “explode” on entry into a mesh with links and photos.

Most virtual environments support user representations, although none are shown in Figure 1. Such representations of actants or users are called “avatars” or “embodiments”. Interactions between users can be achieved through auditory or visual (textual or graphical) communication channels. Interaction between users and the system can be effected through a variety of means (e.g. mouse-driven, speech driven, typed command driven).

In this paper, we briefly discuss a number of recently developed systems and applications. Prior to that, however, we discuss the nature of collaborative and cooperative work activities and discuss the potential benefits of using virtual environments to support such activities. In effect, we address the design requirements for effectively placing the ‘C’ in **C**VE. Finally, we detail a number of future research directions that ultimately will bring together current practice and future possibilities.

2. ADDING THE C TO VE: DESIGNING TO SUPPORT COLLABORATIVE WORK

In addition to their immediate aesthetic appeal, systems and applications are designed to serve a purpose. To be fit for their intended purpose(s), systems must be designed with intended user’s tasks explicitly considered (Newman and Lamming, 1995; Norman, 1988; Rodden *et al*,, 1992). Specifically, in order to support collaborative and cooperative activities, it is important that virtual environments provide task appropriate information representation and communication tools, and appropriately designed landscapes in which activities can occur. In this section of the paper we turn to consideration of some characteristic features of collaborative work to consider how to harness the potential of distributed virtual environments to create effective CVEs.

2.1 Observations of co-operative and collaborative work

When designing systems to support collaborative work we can learn a great deal from observations of people working and collaborating together in conventional settings (Moran and Anderson, 1990). A number of field studies in recent years have offered insights into the nature of collaborative work (Heath and Luff, 1991; Engestr❾m and Middleton, 1996). Although we offer some general points below, we also acknowledge that cooperative and collaborative activities involve considerable negotiation, and teams vary tremendously in their negotiation strategies as well as in their task accomplishment processes (see Lyytinen *et al*, 1994). Thus, the design of a CVE for any task context would require further investigation of *that* task context within a formative design framework.

2.1.1 Characterizing collaborative work: Transitions between shared and individual activities

Work by numerous researchers testifies that in all 'real world' domains, collaborative work involves the interleaving of individual and group effort. Collaborative work thus involves considerable complex information exchange (Bellotti and Rogers, 1997; Harper, 1997; Heath and Luff, 1991; Heath and Luff, 1996; Hutchins and Klausen, 1996; Suchman, 1996). These interleaved, singular-to-shared activities require considerable explicit and tacit communication between collaborators to be successful. Individuals need to negotiate shared understandings of task goals, of task decomposition and sub-task allocation and of task/sub-task progress. It is important that collaborators know what is currently being done and what has been done in the context of the task goals.

To date, most CVEs have been used as meeting places where group activities are the central task (e.g. Greenhalgh, 1997). However, it is easy to imagine situations where more complex, interleaved individual and collaborative activities would be carried out within one VE with team members moving continually between individual and collaborative activities.

2.1.2 Characterizing collaborative work: Flexible and multiple viewpoints

Tasks often require use of multiple representations, each tailored to a different points of view and different subtasks. For example, Bellotti and Rogers (1997) offer a detailed analysis of the production of a daily newspaper. In this process many different representations are used to design the layout of the paper; these vary from hand-drawn to computer generated and reflect different task requirements. In certain cases, one individual may require multiple representations to reflect different aspects of their task(s), whilst in other cases different individuals may require tailored representations to provide information specific to their tasks. For example, consider an architectural design review of a building. Whilst the people taking part in the review would need an overview of the building they might only want to see features relating to their specialty in detail. Thus, an electrician might only want to see wiring plans in detail, but not necessarily see the plans for the plumbing, except in cases where there was a potential conflict. Smith (1996) offers another example. Multi-lingual text displays in virtual environments enables users to explore the same basic environment whilst seeing all textual annotations in their preferred language.

There is also a broader social context within which visualizations may need to vary; if the data being visualized is commercially sensitive, then viewing of certain pieces of information might need to be restricted to a subset of the user population. If the virtual environment does not support subjective views then the users are forced to agree on a common (possibly non-optimal) visualization style.

2.1.3 Characterizing collaborative work: Sharing context

Shared context is crucial for collaborative activities. ‘Shared context’ can mean many things; it can mean shared knowledge of each other’s current activities, shared knowledge of others’ past activities, shared artifacts and shared environment. Together, these lead to shared understandings.

Shared physical space facilitates or ‘affords’ shared understandings. As Gaver (1992) states, “in the everyday world, collaboration is situated within a shared, encompassing space, one which is rich with perceptual information about objects and events that can be explored and manipulated”. Within these shared spaces, focused and unfocussed collaboration is accomplished through alignment towards the focal area of the shared activity, such as a shared document (Heath and Luff, 1996) and through gestures like pointing toward portions of the document for added emphasis and clarity. When artifacts are shared not only do they become the subject of communication between users, but also the medium of communication; as one user manipulates an object, changes to it are visible to others in an externalization of processes of change. For the design of systems to support collaborative work, this means that shared artifacts should be visible and available for local negotiation (Dix, 1994; Heath and Luff, 1996) and often the current focus of attention should be indicated. This can drive subsequent activities.

Where actions are not physically collocated and co-temporal providing shared context is more difficult. Work within CSCW has stressed the importance of tools for providing shared context in asynchronous work contexts (meeting capture, version control and so on); such tools provide activity audits and provide ‘awareness’ of others’ activities which in turn provides a sense of shared context.

2.1.4 Characterizing collaborative work: Awareness of others

There are several ways of conceptualizing ‘awareness’. One view is that reflected above, awareness as knowledge of task related activities. For example, Dourish and Bellotti (1992) state that awareness is an “understanding of the activities of others, which provides a context for your own activity”. This is awareness as an audit of activity, and is the predominant notion of awareness in groupware where activities are often asynchronous and tasks are ‘handed off’ between team members. Such a view of awareness centralises intentional awareness. Awareness has also been discussed in terms of feelings of ‘co-presence’. Notably this sense of awareness intersects with work by Slater and Usoh (1994) on sense of presence, where ‘presence’ refers to a sense of immersion in an environment. Awareness as co-presence involves consideration of peripheral as well as focussed attention and more accurately characterises what occurs when team members are engaged in parallel but independent ongoing activities. In contrast to intentional awareness, such tasks often require moment-to-moment, peripheral co-ordination (Kendon, 1990). For example air traffic controllers are not aware of each other’s moment-to-moment activities but peripheral vision and background sound all provide information such that if a disruption occurs, unplanned collaborative activities can ensue (see also Heath and Luff’s analysis of London Underground workers, Heath and Luff, 1996). Some consideration of providing such moment-to-moment awareness or sense of co-presence exists in work on video portholes and video tunnels where offices are linked with video cameras and monitors (Buxton and Moran, 1990; Gaver, 1992).

Awareness can also relate to activities outside of the current task context where one is interested in the activities of a collaborator who is not currently present and who may not be working on the shared task. Often we need to know where to get hold of someone and/or need to adjust our plans on the basis of when someone will be back. In everyday life, *.plan* files, answerphones and a vacation email messages play this role.

2.1.5 Characterizing collaborative work: Negotiation and communication

Negotiation through talk is important for collaboration (Wardhaugh, 1985). Collaborative work requires the negotiation not only of task related content, but also of task structure in terms roles and activities and task/sub-task allocations.

Conversation analytic studies of negotiation at work have detailed how subtle verbal and nonverbal contribute to such negotiations. Nonverbal cues are crucial for most conversations. Studies of nonverbal behavior suggest facial expression, body posture and gesture carry from 60%-90% of the information in the verbal message being conveyed. Finally, cues from evaluation of dress, posture, behavioral traits and mannerisms provide much framing context for interpretations of verbal negotiations (Hewstone, *et al.*, 1988; Kendon, 1990).

Summary

The issues raised above are relevant for the design of CVEs. For example, ideal modes of communication for particular tasks need to be considered and representation appropriateness has to be taken into account. The design of the virtual space (the landscape and the data representation) is also crucial as this has implications for the ways in which the environment is used and experienced. The capabilities of the user embodiments in terms of appearance, gestural capabilities, facial expressions and so on clearly have implications for the interactions that are possible. In the next section we address some of the issues raised and broadly consider the differences between using CVEs and using more traditional groupware and CSCW products to support collaborative work.

3. CVEs AS A PARADIGM SHIFT IN COMPUTER SUPPORTED COLLABORATIVE WORK (CSCW)

A number of features are central to CVEs. These features distinguish CVEs from CSCW applications and Groupware applications like Lotus Notes (Baecker, 1992; Bannon, 1993; Johansen, 1988; Kraemer and King, 1988) which also provide distributed access to shared contexts. Other examples of groupware are desktop conferencing, video conferencing, co-authoring features and applications, bulletin boards, email, meeting support systems, workflow systems, group calendars, meeting support systems and voice applications. These applications focus on providing team members with access to the same information and a sense of what has been achieved within a collaborative or co-operative task. Although in our discussion above we focused on complex, highly negotiated and interwoven collaborative tasks, some collaborative and co-operative activities are more are routine cases of information exchange or task execution (Bowers, 1992). For these routine activities where clear categories for hand-off are available traditional systems work well. Participants can check into a system, see what has been entered since the last occasion, enter comments and then depart. However, Grinter (1997) has shown how such systems are often designed to incorporate models of work flow; such models sometimes constrain communication rather than facilitate it. There are numerous examples, where such systems do not work effectively (Button and Sharrock, 1994); arguably these occasions are those that require more complex negotiation and interaction, the characteristics which were highlighted in Section 2 above. Specifically, early on in task decomposition and in less well structured tasks like collaborative design activities, co-workers spend more time in negotiations, discussion and task specification. At these stages offering structure prematurely can cause frustration and reduce creativity.

We believe that CVEs can support many forms of communication both to support routinised tasks but also negotiated, interwoven tasks for which a more fine-grained collaboration is ideal. In the next section we will concentrate on those collaborative activities in considering how CVEs can complement existing systems by providing support for tasks that are not well supported by CSCW and Groupware. The CVEs we focus on in this paper are distinctive in their use of sophisticated graphics (revealing their roots in virtual reality technologies), in their use of embodiments and in the way shared space is conceptualized. The last two points are shared with technologically less sophisticated CVEs like MUDs and MOOS (see papers by Cicognani and Hand *et al* in this volume).

3.1 CVEs and the concept of space

CVEs represent a paradigm shift in that they provide a space that contains or encompasses data representations *and* users. To illustrate, Grudin and Poltrock in their 1996 CHI tutorial on CSCW systems, offer three central metaphors that have accompanied the development of computer use. The first metaphor is the computer as a tool that literally computes, where program and data are fed into computing machinery on cards and tape. The second metaphor is the computer as container, where the programs and data are held in the machine (often displayed on a desktop) and files can be ‘downloaded into’ the computer, by ftp for example. Their third metaphor, and the predominant metaphor in CSCW applications, is the computer as window, where we can “view and converse with distant people and objects”.

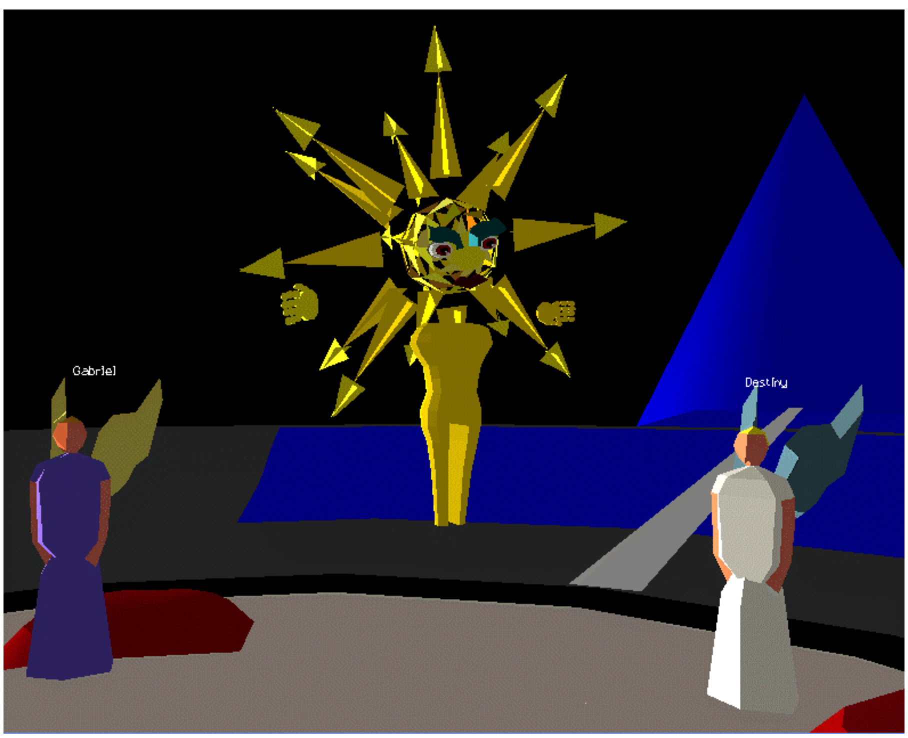

Figure 2: embodiments in a virtual space

However, rather than a window, CVEs represent the computer as a malleable *space*, a space *in which* to build and utilize shared places for work and leisure. CVEs provide a terrain or digital landscape that can be 'inhabited' or 'populated' by individuals and data, encouraging a sense of shared space or place. Users, in the form of embodiments or avatars (see Figure 2), are free to navigate *through* the space, encountering each other, artifacts and data objects and are free to communicate with each using verbal and non-verbal communication through visual and auditory channels.

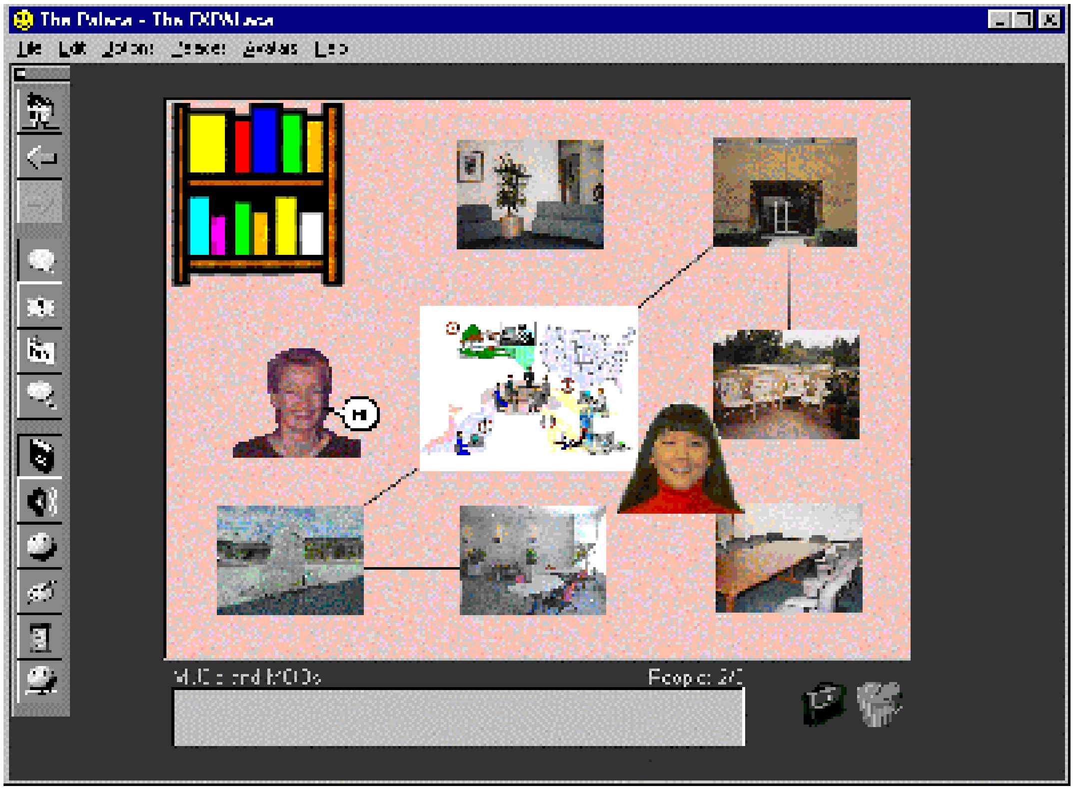

Figure 3: A room created using The Palace with two avatar embodiments

As stated above, although conceptually simpler and technologically much lighter weight, this notion of providing places for interaction is also reflected within environments like MUDs and MOOs despite the lack of sophisticated graphics (Curtis and Nichols, 1992; see Hand *et al*, this volume; Tennison and Churchill, 1995; Tromp, 1993). This is carried into simple graphical MUDs and MOOs (e.g. The Palace) which offer limited avatar or embodiment representations which are generally limited, acting more as "placeholders" than active, gesturing embodiments. Figure 4 illustrates this, showing a snapshot of a "room" designed in The Palace; the avatars shown in the figure are considerably less sophisticated than those shown in Figure 3 and space here is 2D rather than 3D as shown above. Although in this paper, we concentrate on collaboration within 3D graphical environments, studies indicate that users do indeed feel a sense of space and of co-presence with others in MUD and MOO rooms, despite the lack of sophisticated graphics (Adams *et al*, 1997; Tennison and Churchill, 1996; Tromp, 1993).

3.2 CVEs and collaborative work

Having laid out this conceptual ground, we shall now consider how 3D graphical CVEs can support collaborative work by considering some of the issues raised above. The possibilities for flexibility in representation specification and in the establishment of shared virtual spaces means that CVEs are able to support many of the activities outlined above.

It is clear that CVEs should allow users to switch smoothly between individual and collaborative tasks. This means being able to alter representations and navigate easily, as well as being able to communicate quickly and efficiently regarding current activities. Currently many systems are rather clumsy in the way interaction and modality shifts are handled, and the tools for navigation do not always provide the easiest mapping from 'real world' to VE movement, but research developments are improving this situation. For example, investigations into the use of speech input for altering views (see Bersot *et al*., this volume) and the use of a bicycle as an input device in Diamond Park (Waters *et al*., 1996).

Sharing a space with others can facilitate communication and collaboration. It has been argued that a shared virtual space that adheres to shared conventions (in this case real world spatial conventions although this need not necessarily be the case) thus provides a natural and intuitive way to navigate and to interact with individuals and groups (Benford *et al*., 1994). Certainly, research from many disciplines confirms that the design of spaces affects and is affected by human social action, social interaction, perception and cognition (Barwise and Hammer, 1995; Freska, 1991; Giddens, 1984; Haarslev, 1995; Werlen, 1993). A potentially infinite virtual space provides potentially infinite opportunities for individuals to define and delimit regions for particular activities. Such collaboratively defined, bordered places within the virtual space provide the possibility of *framing* social behaviours and social interactions (Goffman, 1972). Experiences in virtual communities like AlphaWorld (and De Digitale Stad (http://www.dds.nl/) confirm this view (Damer, 1997); within these digital spaces there exist numerous places where like-minded indivdiuals gather and interact.

The flexibility of virtual environments thus supports flexible, negotiated interactions. By contrast, traditional CSCW applications and more spatially oriented applications like shared drawing tools (e.g. VideoDraw (Tang and Minneman, 1990) and Clearboard (Ishii and Kobayashi, 1992)) offer a shared window onto a shared task, usually on a point-to-point basis. They offer less support for the flexible and spontaneous altering of shared frames of reference or the redrawing of shared places and regions in which renegotiated activities can take place. Similarly, the negotiation of multiple places within the VE for multiple groups to collaborate in parallel is not as easily established. As such, along with video conferencing and video awareness tools, these offer what has been called "discontinuous and arbitrary space" (Gaver, 1992). Virtual environments potentially provide many places within one encompassing space or place where many groups can define regions of activity or one group can define regions for different activities, all within one continuous space which adheres to shared conventions drawn from the physical world.

Further, the flexibility of virtual environments enables flexibility in the design of shared data representations. It is clear from the discussion above that tasks often require many levels of expertise and different levels of information. Although most current multi-user virtual reality systems provide a highly 'objective' virtual environment where all users see the environment in the same way (albeit from different viewpoints) and all users see the same objects in the same places with the same appearances, this need not be the case. The current predominance of the objective view is partly due to the fact that multi-user VR systems evolved from single-user systems that have been extended to support a number of users. This aspect of the evolution of VR systems is similar to the way that groupware systems evolved from systems such as Shared X [Gust, 1988], that simply replicated an application's user interface to multiple users. However, as discussed above, visualizations of abstract information are not fixed; the choice of visualization style can vary widely since the source data has no intrinsic appearance. Given this freedom of choice it is likely that users will form their own preferences for the display of particular datasets for particular tasks. However, if the virtual environment is capable of supporting subjective views then users are free to choose their own preferred visualization style. Subjective or differing views may also be appropriate for other cognitive reasons. These requirements can be accommodated with CVEs if each user has an independently controlled viewpoint from which they can inspect the virtual environment. This corresponds to our experience of the real world and as such we expect people to be skilled at relating to viewpoints and reasoning about these kinds of subjective effects. One way of achieving this individuality in representation whilst maintaining an easily accessible shared context is through level-of-detail (LOD) effects. LOD effects represent a common technique in interactive computer graphics and act to provide several different representations of a single artifact each with a different level of complexity. When viewing the object from a distance the user is shown a simple version of the artifact which is replaced by more detailed representations as the user approaches.

Similarly, flexibility in the design of embodiments means that embodiments can be tailored to task requirements, particularly the degree of interaction between users. Within CVEs, auditory and visual channels support real-time conversation and within 3D graphical CVEs embodiments support the use of gesture and non-verbal cues to support interactions. As discussed above, gesture and facial play an important role in human interaction. Benford et. al. (1997a) state that "the inhabitants of collaborative virtual environments (and other kinds of collaborative systems) ought to be visible to themselves and to others through a process of direct and sufficiently rich embodiment". Simply, individuated, gesturing embodiments combined with good information channels (visual and auditory) enable us to know who's who, where they are and what they're doing. Users have to recognize who someone is from their embodiment(s), and be able to differentiate between autonomous software agents for example and 'avatars' in the true sense (embodiments behind which are users in 'live' interaction). Embodiments provide a sense of location within the shared space, acting as 'placeholders' as mentioned above, but also providing dynamic cues as to current foci and activities. Thus, embodiments within a spatial framework provide a sense of awareness of others through cues akin to those in the physical world. For example, they offer proximity cues through use of auditory and visual cues, activity cues through their movements and informational cues through gesture, speech and intonation.

In contrast to video technologies that offer 'a talking heads' vision of one's collaborators (ideally reciprocal), embodiments within CVEs offer greater scope for use of body movements and spatial orientation. These offer complex cues for interaction and for development of shared contexts; such cues are critical in our everyday interactions in the 'real' world when we are physically collocated with others (Kendon, 1990). Collaborators can, for example, point to objects or to portions of objects and move around objects to have multiple shared viewpoints. In many CSCW systems this process becomes more cumbersome because users have to point to objects via input devices and using complex sentences, but this issue is being addressed in current research. Notably, this capability is not available within telephone conferencing or within textual MUDs and MOOs. Graphical MOOs provide only limited possibilities for this (Adams *et al*,, 1997).

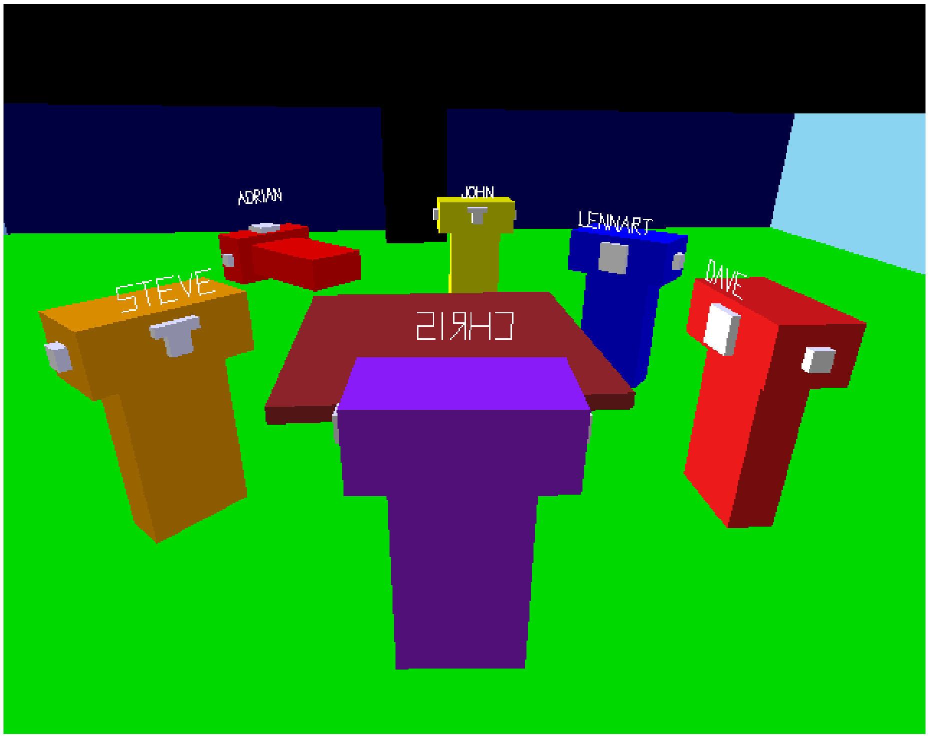

Figure 4: Embodiments in MASSIVE-1

Currently facial expressions and fine grained limb movements are not well developed for CVE embodiments. However, we believe that tailorable representations will be offered in the future; these will have possibilities for more individuation, greater scope for self representation and increasingly fine-tuned non verbal behaviours. Even at the current level of embodiment design however, it is possible to use conventions to signify certain things. For example, embodiments can convey information about the availability for interaction. This is illustrated in Figure 4. Here several embodiments are shown, positioned around a table. One embodiment is ‘lying down’; within this group, this position is a convention that indicates that the person behind the embodiment is not present. The embodiment owner is engaged elsewhere. This allows the other group members to know there will be no input from that person. Similarly, the ‘faces’ of the embodiments indicate the communication capabilities of different team members. The embodiment with a ‘T’ indicates this user has only text and not audio chat capabilities.

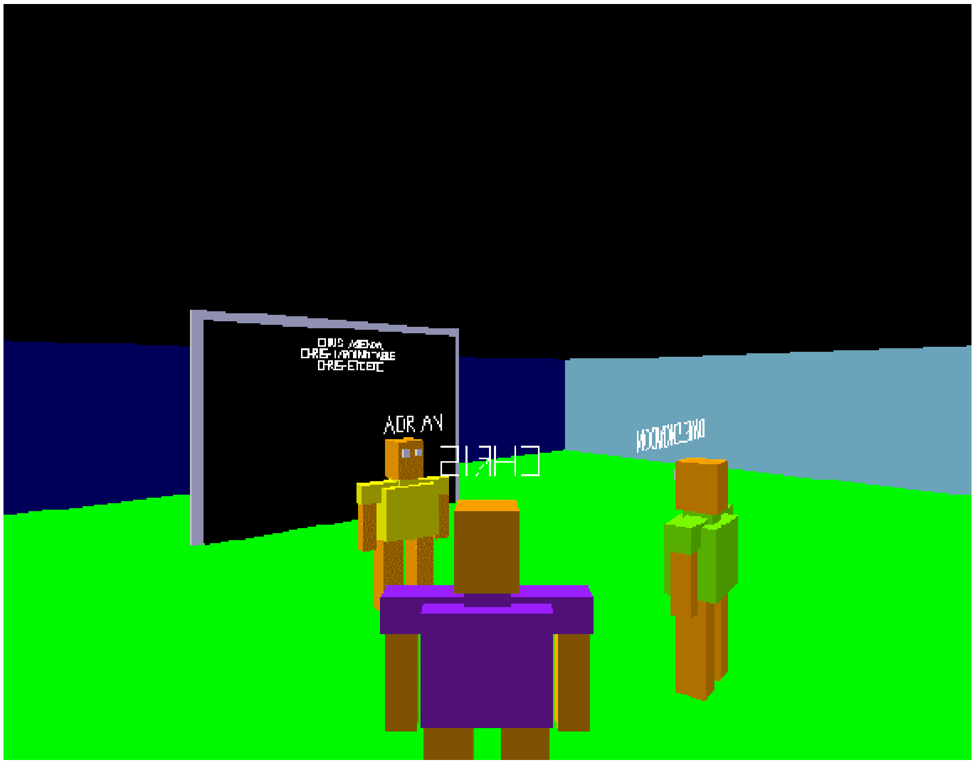

Figure 5: Embodiments in MASSIVE-1 around a shared blackboard

Embodiments can also support historical awareness of who has been present and what activities they performed. Semi-autonomous embodiments can act on a user's behalf, offering useful information when the owner/author/user is not currently present (Greenhalgh and Benford, 1996). A user can also leave many embodiments around the VE; thus people can be in several places in one CVE, or in several CVEs at the same time. Users can also support asynchronous interactions by leaving each other messages within the VE. In Figure 5, several embodiments are shown standing around a virtual blackboard on which a message is etched. Similarly in Laurel and Strickland's (1994) art installation *Placeholder*, visitors were able to leave voicemail for later review by themselves or others. These examples show the capabilities for supporting asynchronous activities as well as synchronous activities.

4. EXAMPLE SYSTEMS

In this section we shall briefly describe a small number of existing systems and applications, each chosen to highlight a particular point of interest and illustrate the range of applications for which CVEs are being employed which include. These include information visualization (VR-VIBE), tele-conferencing (MASSIVE-1), battlefield simulation (NPSNET) and social and artistic events (MASSIVE / MASSIVE-2 &

Diamond Park / SPLINE). For a more complete review of existing VE systems the reader is directed to Snowdon *et al* (1996).

4.1 DIVE

DIVE (Hagsand, 1996) is a multi-user virtual environment developed at the Swedish Institute of computer science. DIVE is one of the more mature multi-user virtual environments and is being used worldwide for developing VE applications. DIVE 3 supports scripting using Tcl/Tk (Ousterhout, 1994). This is a powerful feature and allows DIVE objects to have associated behaviour without the need for a heavyweight process to control it. Another more interesting feature is the capability for one DIVE process to send a Tcl/Tk script to be executed in another DIVE process. This allows, for example, a DIVE application to send a script to the DIVE user client which extends the user interface to add application specific controls.

A number of applications have been constructed using DIVE, one of these is VR-VIBE (Benford *et al*., 1995) which provides a multi-user visualization of bibliographic data and allows users to browse and search the corpus of documents. Developments reflected in DIVE 3 discussed above are exploited by the latest version of VR-VIBE, which provides a variety of visualizations and user-interface styles. The main application process is only concerned with providing services for managing a visualization and all the user interface components are entirely separate and written in Tcl/Tk. It is therefore possible to completely replace the VR-VIBE interface without changing the main application at all.

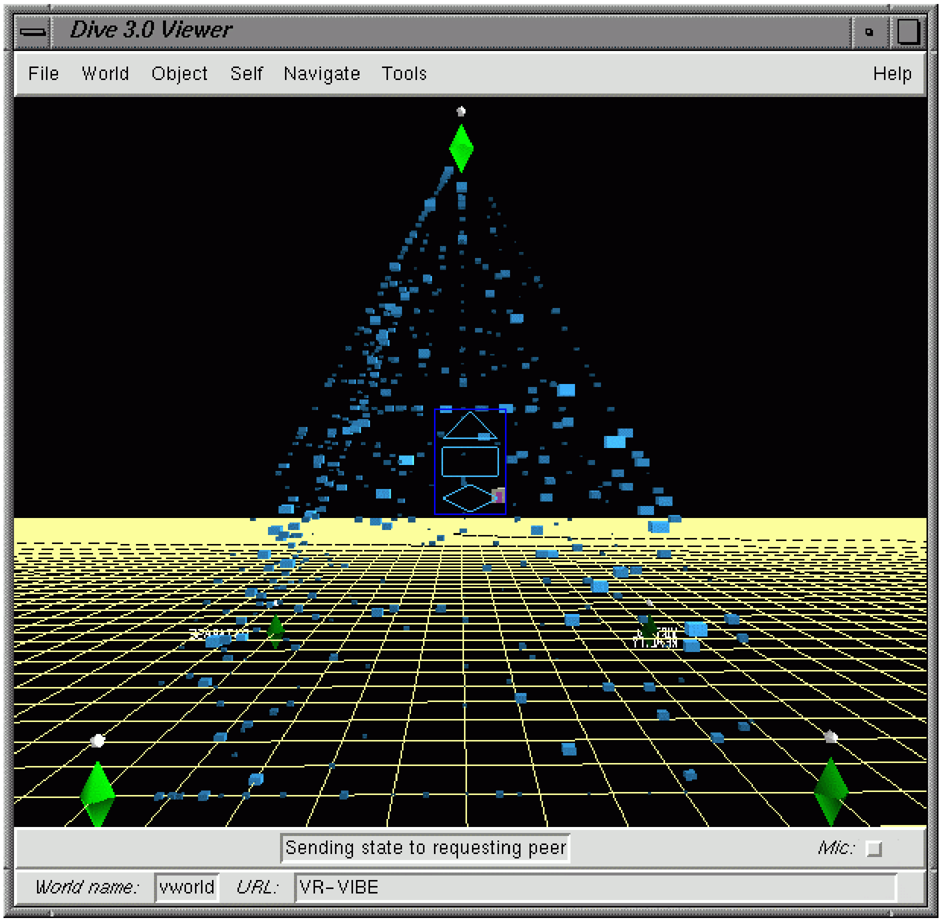

Figure 6: VR-VIBE visualization of items in a bibliographic database

Figure 6 shows a VR-VIBE 3D representation of items in a database. The green octahedrons (shown at the top centre and bottom left and right of the figure) represent points of interest that have been specified by a user. The blue items in between represent documents in the database. User can navigate through the space of these document icons, and can view the documents by clicking on an icon.

User studies of VR-VIBE have shown promising results; users are able to issue complex queries, navigate successfully through the space and select appropriate items (Churchill *et al*., 1997). Although to date all published work has concerned single user search and browse tasks, current work centres on collaborative search and browse tasks using VR-VIBE visualizations as shared, distributed environments.

4.2 MASSIVE

MASSIVE-1 (Greenhalgh and Benford., 1995) is a distributed VR system targeted initially at tele-conferencing applications. MASSIVE currently runs on Sun and Silicon Graphics workstations. The main feature of MASSIVE-1 is that it supports multiple users, applications and worlds that are connected by portals. A noteworthy feature of MASSIVE is that a user of a powerful workstation running graphical, audio and text interfaces can interact, albeit in a limited way, with another user who has only a vt100 running a text interface. The text user is represented in the graphical medium and vice versa. The users can communicate by emoting, by typed text, and the text adapters (text to speech and text to graphics) allow graphical users to follow the activities of text-only users. For users with more technological capabilities, it is possible to employ an arbitrary combination of graphics, text and audio interfaces.

The COMIC spatial model of interaction (Benford & Fahl②n, 1993) controls all interactions. This model is intended to support interaction and cooperation in heavily populated virtual spaces. To this end, MASSIVE-1 concentrates on facilitating user-user interaction, while user-object interaction is relatively undeveloped. This spatial model of interaction uses the concepts of aura, focus and nimbus (Benford *et al.*, 1997). Users can explicitly control their interactions with other objects; specifically, users are able to control how aware they are of other objects or users, and how aware those objects or users are of them. Given the discussion above on the importance of communication, feelings of co-presence and awareness for collaborative activities, this is an important feature.

MASSIVE currently supports up to about 10 mutually aware users in a given world. It has been used over the LAN for internal research group meetings, and has been used over the UK's SuperJANET network to hold three site meetings between Nottingham, Lancaster and QMW in London and over the internet to hold five-site meetings between Nottingham, Lancaster, Stockholm (2 sites) and Bonn.

MASSIVE-2 combines ideas from MASSIVE-1 (Greenhalgh and Benford., 1995) and AVIARY (Snowdon and West, 1994). From MASSIVE-1 is borrowed the use of the COMIC spatial model of interaction. From AVIARY is borrowed the architecture based on a distributed object system. MASSIVE-2 also adds some novel functionality. This includes hierarchical compositions of regions in virtual worlds which can be mapped to separated multicast groups, different media which can either share or use separate multicast groups, support for the new extended spatial model of interaction which includes third parties (Benford *et al.*, 1997a), regions and abstractions and an extensible object oriented (class-based) developers API.

A number of evaluations of the MASSIVE distributed environment have been conducted. Early evaluation studies concentrated on group collaborative and competitive games. These indicated that users were able to easily manipulate aspects of the spatial model to control awareness of others and to vary levels of privacy in conversations. Having explicit control over these levels of interaction proved much less frustrating for users than when no such control was provided. In the everyday world, we have control over the level of privacy of any utterance by whispering or shouting. This expectation was clear amongst our users (Churchill *et al.,* Mann, 1996). The studies also indicated a need to support clear, moment-to-moment communication between team members and to consider what information needed to be shared and how quickly, reinforcing the user centred design perspective outlined above.

A more extensive but less formal evaluation was the MASSIVE VR/poetry performance (Benford *et al,* 1997c) which took place as part of the Nottingham NOW ninety6 arts festival. This event was an experiment in using a CVE as a performance space. Four rap poets performed (in sequence) on a stage to a live audience. Position trackers were attached to the poets' heads and hands and used to control the embodiment representing the poet in a virtual space. In a separate room 10 Silicon Graphics workstations running the MASSIVE-2 system allowed members of the public to see and hear the virtual representation of the poet, to explore the virtual environment and to talk with one another via a real-time audio connection. Views from the virtual environment were projected onto screens in both rooms so that people not currently logged on could follow the activity in the virtual environment. The results indicated that users enjoyed the event very much and were able to easily navigate the fairly extensive virtual environment. Communication was also easy. Navigation and communication were so easy in fact that the virtual audience members spent more time communicating with each other and navigating the space than they did watching the poetry event. The poets themselves were not aware of the activities of the virtual audience members, as they were engaged in performing to the ëliveí audience.

A number of design recommendations deriving from this event are detailed by Benford *et al* (1997). These mainly concern the level of control embodiment are granted regarding the use of the spatial model of interaction described above. Notably the recommendations offered are pertinent to the specific type of event considered, in this case a public performance with a stage and with virtual audience members. In the context discussed above, the authors recommend that different embodiment capabilities depending on the activity to be supported. Smaller foci and nimbus when an event needs to be the centre of attention and larger foci and nimbus during exploration. They also discuss the possibility of using an object centred navigation where central objects could be designated as focal points about which embodiment can move. The last point is particularly interesting: the authors argue that ëlockingí embodiments to an object may enhance the development of a mutual awareness between team members.

4.3 NPSNET

NPSNET (Zyda, 1993) is a networked VR system designed for military training and simulation with the goal of supporting large numbers of participants. NPSNET is a single application rather than a general purpose system intended to support applications. Its eventual goal is to support hundreds of users. The means by which it attempts to do this are therefore of interest to the designers of large scale multi-user VR systems. NPSNET uses the standard DIS (Distributed Interactive Simulation) protocol to communicate between entities taking part in the simulation (IEEE, 1993). IP- MULTICAST is used to transmit updates to reduce bandwidth requirements.

4.4 SPLINE

SPLINE (Waters, *et al*., 1996) is a CVE system that supports both audio and video interaction between users and is designed to support large numbers of users. The principle application for SPLINE so far has been to create Diamond Park, a social virtual environment containing a square mile of terrain. In addition to a desktop workstation based interface, Diamond Park supports a novel user interface based on modified recumbent exercise bicycles which allows users to cycle around the environment. To date no formal user testing has been done of Diamond Park or other applications based in SPLINE (W. Yerazunis, Mitsubishi Electric, personal communication).

5. SUMMARY AND FUTURE WORK

In this paper we have outlined how 3D virtual environments can provide support for co-temporal and asynchronous collaborative activities. CVEs can provide support for synchronous activities, unlike email and bulletin boards, and can provide real-time support for the sharing of visual artifacts, unlike telephone conferencing facilities. CVEs can also support asynchronous activities: leaving information within a virtual environment for later review is trivial.

In terms of architecture, CVEs provide a single space within which data sets, individuals and frames of reference are rendered, ideally using a compatible set of methods, within a shared technological framework. This offers an advantage of flexibility; in principle, any digital data stream can be rendered within a CVE. Thus, it is possible to provide video conferencing facilities by placing a video stream onto an object within the virtual environment itself (Benford and Fahl②n, 1993; Reynard and Benford, 1997). This flexibility means a CVE may offer shared environment in which one can video conference whilst sharing multiple abstracted visualizations with virtually co-present team members. Further, as all entities exist within one arena, they can all be designed to adhere to the same underlying spatial model, in the way that items in the ‘real’ world are constrained by the same environmental constraints. By contrast, to date media spaces (Bly *et al*, 1993) tend to offer assemblies of various modalities for interaction, still largely without conscious effort to integrate these communication channels into coherent spatial metaphors which reflect the affordances of continuous or what is considered ‘real’ space (Gaver, 1992).

In addition, groups vary in size. Nearly all real-time CSCW systems (e.g. multi-media conferencing and media spaces) have focused on small scale interaction between a few concurrent users. In order to enable synchronous communication between large numbers of users the underlying approach must have some inherent property that is scaleable. CVEs are inherently scaleable. Since environment designers are free to use as much 3D space as is desired, CVEs have the potential to accommodate large numbers of users. However, there are significant technological challenges which must first be overcome if CVEs are to support hundreds of thousands of users and still provide real-time audio and rich content whilst providing acceptable levels of performance.

At the usage levels, perceptual, cognitive and social issues are increasingly of concern in the design of useable and useful CVEs. Evaluation studies and work within CSCW continue to consider the nature of collaborative and co-operative activities and these results are feeding into the design of systems (see Billinghurst, this volume; Bolzoni *et al*, 1996). These results combined with technological developments are offering greater possibilities for the increased use of virtual environments to support collaborative and co-operative activities in the future.

Acknowledgments

The authors would like to thank Rob Macredie for his encouragement to produce this paper and for encouraging the publication of the papers in this special issue. The authors would also like to thank all the participants of CVE'96 in Nottingham and members of Nottingham University's Communications Research Group.

References

Adams, L., Toomey, L. and Churchill, E.F. (1997) Meetings in a Virtual Space: Creating a Digital Document. Paper accepted for presentation at HICCSí98, Hawaii, January 1998.

Baecker, R. M. (1993) *Readings in Groupware and Computer Supported Cooperative Work*. Morgan Kaufman.

Bannon, L. J. (1993) CSCW: An Initial Exploration. In *Scandinavian Journal of Information Systems, Vol. 5*, pp 3-24, 1993.

Barwise, J. and Hammer, E. (1995) Diagrams and the Concept of Logical System.. In D. Gabbay (Ed) *What is a Logical System?* UK: Oxford University Press.

Bellotti, V. and Rogers, Y. (1997) From Web Press to Web Pressure: Multimedia Representations and Multimedia Publishing. In *Proceedings CHI'97, ACM Conference on Human Factors in Computing,* Atlanta Georgia, USA, March 1997.

Benford, S. D., Bowers, J. M., Fahl②n, L. E., Greenhalgh, C. M., Snowdon, D. N. (1997a) Embodiments, Avatars, Clones and Agents for Multi-user, Multi-sensory Virtual Worlds. In *Multimedia Systems, 5, (2),* Springer-Verlag 1997, pg 93-104.

Benford, S., Greenhalgh, C., Snowdon, D., and Bullock, A. (1997b) Staging a Public Poetry Performance in a Collaborative Virtual Environment. *In Proceedings of the 5th European Conference on Computer Supported Cooperative Work (ECSCW'97)* September, 1997, Lancaster, UK Kluwer Academic Publishers.

Benford, S., Snowdon, D., Brown, C., Reynard, G. and Ingram, R (1997c) The populated web: Browsing, Searching and Inhabiting the WWW Using Collaborative Virtual Environments. *In Proceedings of Interact'97 - the 6th IFIP Conference on Human-Computer Interaction incorporating OZCHI97 and APCHI97*, July, 1997, Sydney, Australia.

Benford, S., Snowdon, D., Greenhalgh, C., Ingram, R., and Knox, I., (1995) VR-VIBE: A Virtual Environment for Co-operative Information Retrieval. In *Computer Graphics Forum (Proceedings Eurographics'95), 14(2),* Post and G❾bel eds., Blackwell, 1995, pg 349-360.

Benford, S. D., Bowers, J. M., Fahlen, L. E., Mariani, J. and Rodden, T. R.(1994) Supporting Co-operative Work in Virtual Environments. In *The Computer Journal, 37 (8)*, Oxford University Press, 1994.

Benford, S. D., and Fahl②n, L. (1993), A spatial model of interaction in virtual environments, *In Proceedings of Third European Conference on Computer Supported Cooperative Work (ECSCW'93),* Milan, Italy, September 1993.

Bly, Sara A., Harrison, S.R. and Irwin, Susan (1993) Media Spaces: Bringing People Together in a Video, Audio and Computing Environment. *Communications of the ACM 36* (1) 193: 28-47.

Bolzoni, M., Riva, G., and Melis, L. Interacting with people and objects in virtual environments. Metaphors for User Centred Design. In *Proceedings of CVE 96 Workshop on Collaborative Virtual Environments*, 19-20 September, 1996.

Bowers, J. (1992) The Politics of Formalism. In M. Lea (Ed). Contexts of Computer Mediated Communication. Harvester Wheatsheaf, New York, 1992, pp 232-261.

Button, G. and Sharrock, W. (1995) Occasioned Practices in the Work of Software Engineers. In J. Goguen and M. Jirotka (Eds) *Requirements Engineering*. Academic Press Ltd. London, UK, 1994, pp 217-240

Buxton, W. and Moran, T. (1990) EuroPARCís integrated interactive intermedia facility (iiif): Early experiences. In *Proceedings of the IFIP WG8.4 Conference on Multi-User Interfaces and Applications, Herakleion*, Crete, September 1990.

Churchill, E.F., Benford, S. and Mann, S. (1996) Collaboration in Virtual Enviornments. Internal Technical Report, Cognitive Ergonomics Group, Department of Psychology, University of Nottingham, Nottingham, UK.

Churchill, E.F., Snowdon, D. Benford, S. and Dhanda, P. (1997) Using VR-VIBE; browsing and searching for documents in 3D space. In Smith, M.J., Salvendy, G. and Koubek, R.J. (Eds) *Advances in Human Factors/Ergonomics 21B; Design of Computing Systems: Social and Ergonomic Considerations. Proceedings of HCI International*., San Francisco, August 1997.

Curtis, P. and Nichols, D. (1992). MUDs Grow Up: Social Virtual Reality in the Real World, *Proceedings of the Third International Conference on Cyberspace*, Austin, TX, 1993.

Damer, B. (1997) Interacting and Designing in Virtual Worlds on the Internet. BayCHI Tutorial, July 31st, 1997.

Dix, A. (1994) Computer Supported Cooperative Work: A framework. In D. Rosenberg and C. Hutchinson (Eds) *Design Issues in CSCW*. Springer-Verlag, Berlin.

Dourish, P. and Bellotti, V. (1992) Awareness and Collaboration in Shared Workspaces. In *Proceedings of CSCW'92*, November, 1992: ACM Press.

Dyson, E. (1992) A Framework for Groupware. In *Proceedings of Groupware`92*, Morgan Kaufman 1992, pg 10-20.

Engestr❾m, Y. and Middleton, D. (1996) *Cognition and Communication at Work*. Cambridge University Press, New York.

Freska, C. (1991) Qualitative Spatial Reasoning. In D. M. Mark and A.U. Frank. *Cognitive and Linguistic Aspects of Geographic Space*, NATO ASI Series D. pg 361-372.

Gaver, W.W. (1992) The Affordances of Media Spaces for Collaboration. In *Proceedings of CSCW 92,* November 1992.

Giddens, A. (1984) *The Constitution of Society*. Polity, Cambridge.

Goffman, E. (1972*) Interaction Ritual*. Allen Lane, London.

Greenhalgh C. (1997) Analyzing movement and world transition in virtual reality teleconferencing. . In *Proceedings of 5th European Conference on Computer Supported Cooperative Work (ECSCW'97)* September, 1997, Lancaster, UK Kluwer Academic Publishers.

Greenhalgh C. and Benford, S. (1995) MASSIVE: A Virtual Reality System for Tele-conferencing. In *ACM Transactions on Computer Human Interfaces (TOCHI), 2(3),* ACM Press September 1995, pg 239-261.

Grinter, R. (1997) Doing Software Development: Occasions for Automation and Formalisation. In *Proceedings of ECSCW'97*, Lancaster, September 1997, Kluwer: Dordrecht.

Grudin, J. and Poltrock, S.E. (1996) CSCW, Groupware asnd Workflow: Experiences, State-of-the art and future trends. CHI 96 One Day tutorial, CHI 96, Vancouver, BC, April 14, 1996.

Gust, P. (1988) Shared X: X in a distributed group work environment, 2nd annual X conference, 1988, MIT, Boston, USA

Haarslev, V. (1995) Formal Semanctics of Visual Languages Using Spatial Reasoning. In V. Haarslev (Ed) *Proceedings of 11th IEEE Symposium on Visual Language*s, pg 156-163.

Harper, R. (1997) Gatherers of Information: The Mission Process at theInternational Monetary Fund. In *Proceedings of 5th European Conference on Computer Supported Cooperative Work (ECSCW'97)* September, 1997, Lancaster, UK Kluwer Academic Publishers.

Hagsand, O. (1996) Interactive Multi-user VEs in the DIVE System, IEEE Multimedia, 3(1), pg 30-39.

Heath, C. and Luff, P. (1991) Collaborative activity and technological design: task coordination in London Underground control rooms. In L. Bannon, M. Robinson and K. Schmidt (Eds), Proceedings of ECSCW'91, Kluwer: Dordrecht.

Heath, C. and Luff, P. (1996) Convergent activities: Line control and passenger information on the London Underground. In *Engestrom, Y. and Middleton, D. (1996) Cognition and Communication at Work.* Cambridge University Press, New York.

Hewstone, M., Stroebe, W., Codol, J. and Stephenson, G.M. (1988) *Introduction to Social Psychology*. Basil Blackwell, Oxford, UK.

Hutchins, E. (1990) The technology of team navigation. In J. Galegher, R. Kraut, and C. Egido (Eds) *Intellectual Teamwork: Social and Technical Bases of Collaborative Work*. Hillsdale, NJ: Lawrence Erlbaum Associates

Hutchins, E. and Klausen, T. (1996) Distributed cognition in an airline cockpit. In Engestrom, Y. and Middleton, D. (1996) Cognition and Communication at Work. Cambridge University Press, New York.

Institute of Electrical and Electronic Engineers, (1993), Standard for Information Technology Protocols for Distributed Interactive Simulation, *International Standard ANSI/IEEE* 1278-1993.

Ishii, H. and Kobayashi, M. (1992) Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *Proceedings of CHI'92, Monterey, California, May 3-7, 1992*. ACM: New York, pp 525-532

Johansen, R. (1988) *Groupware: computer support for business teams*. Free Press, New York.

Kraemer, K.L. and King, J.L. (1988) Computer-Based Systems for Cooperative Work and Group Decision Making. *ACM Computing Surveys, Vol 20*, no. 2, pp115-146, June 1988.

Kendon, A. (1990) Behavioural Foundations for the Process of Frame Attunement in Face-to-Face Interaction. In Ginsburg, G.P., Brennan, M.and von Cranach, M. (Eds) *Conducting Interaction*. European Monographs in Social Psychology 35, Cambrdige University Press.

Laurel, B. and Strickland, R, (1994) Placeholder. http://www.wmin.ac.uk/media/VD/Placeholder.html

Lotus, 1997, "Lotus Notes", http://www2.lotus.com/notes.nsf, 1997.

Lyytinen, K., Maaranen, P. and Knuuttila, J. (1994) Groups are not always the same. *In Computer-Supported Co-operative Work (CSCW) 2*: 261-284.

Middleton, D. (1996) Talking Work: Argument, common knowledge, and improvisation in teamwork. In Engestrom, Y. and Middleton, D. (1996) Cognition and Communication at Work. Cambridge University Press, New York.

Moran, T. and Anderson, R.J. (1990) The workaday world as a paradigm for CSCW design. In *Proceedings of CSCW'90*, Los Angeles, California, October 1990,ACM: New York.

Newman, W.M. and Lamming, M.G. (1995) *Interactive System Design*. Addison Wesley Publishing, Wokingham, England.

Norman, D.A. (1988) *The Psychology of Everyday Things*. Basic Books, New York.

Ousterhout, J.K. (1994) Tcl and the Tk Toolkit, Addison-Wesley.

Reynard, G. and Benford, S. (1996) Vivid: A symbiosis between virtual reality and video conferencing, UKERNA Video Conferencing workshop, May 1996, Nottingham University, pages 101-113.

Rodden, T., Mariani, J. and Blair, G. (1992) Supporting Cooperative Applications, International Journal of Computer Supported Cooperative Work (CSCW), 1(1-2), Kluwer 1992.

Slater, M., Usoh, M. and Steed, A. (1994) Depth of Presence in Virtual Environments. In *Presence 3(2*), MIT Press 1994, pg 130-144.

Smith G. (1996) Cooperative Virtual Environments: lessons from 2D multi user interfaces. In *Proceedings of Computer Supported Collaborative Work 1996 (CSCW '96).* Boston, Massachusetts, USA. November 16-20, 1996, Pages 390-398.

Snowdon, D., Greenhalgh, C., Benford, S., Bullock, A., and Brown, (1996) C., A Review of Distributed Architectures for Networked Virtual Reality. *In Virtual Reality: Research, Development and Applications, 2(1)*, Virtual Press.

Snowdon, D., Greenhalgh, C. and Benford, S. (1995) What You See is Not What I See: Subjectivity in Virtual Environments. In *Frameworks for Immersive Virtual Environments (FIVE'95) 18-19th December*, 1995, QMW University of London, UK

Snowdon, D. and West, A.J. (1994) AVIARY: Design issues for future large-scale Virtual Environments. In *MIT Presence*, 3(4), 1994.

Suchman, L. (1996) Constituting shared workspaces. In Engestr❾m, Y. and Middleton, D. (1996*) Cognition and Communication at Work*. Cambridge University Press, New York.

Tang, J. and Minneman, S.L. (1990) VideoDraw: A video interface for collaborative drawing. *In Proceedings of CHI'90, Seattle*, Washington, April 1-5, 1990. ACM: New York, pp 313-200.

Tennison, J. and Churchill, E.F. (1996) Individual and Collaborative Information Retrieval in Virtual Environments. Web site: http://www.psyc.nott.ac.uk/aigr/papers/IR_in_VEs/IR_in_VEs.html

The Palace Inc. (1997), Virtual World Chat Software, http://www.thepalace.com

Tromp, J.G. (1993) Results of two surveys about Spatial Perception and Navigation of a Text-Based Spatial Interface, [WWW document]. URL http://www.cms.dmu.ac.uk/~cph/VR/JolaPaper/jola.html

Wardhaugh, R. (1985) *How Conversation Works*. Basil Blackwell, USA.

Waters, R.C., Anderson, D.B., Barrus, J.W., Brogan, D.C., Casey, M.A., McKeown, S.G., Nitta, T., Sterns, I.B., Yerazunis, W.S. (1996) Diamond Park and Spline: A Social Virtual Reality System with 3D Animation, Spoken Interaction, and Runtime Modifiability. In *Presence: Teleoperators and Virtual Environments, 6(4),* August 1997, MIT Press (MERL technical report TR96-02a, 1996,

Werlen, B. (1993*) Society, Action and Space. An Alternative Human Geography*. Routledge: London and New York.

Wexelblat, A. (1993) The Reality of Cooperation: Virtual Reality and CSCW. In Wexelblat, A. (Ed) *Virtual Reality: Applications and Explorations*, Academic Press, 1993, pg 23-44.

Zyda, M.J., Pratt, D.R., Falby, J.S., Lombardo, C. and Kelleher, K.M. (1993) The Software Required for Computer Generation of Virtual Environments. In *Presence, 2(2*), Spring 1993, pp 130-140